

Autonomous Surface Vehicle 3D Seafloor Reconstruction from Monocular Images and Sonar Data

Eduardo Iscar¹ and Matthew Johnson-Roberson¹

Abstract—Traditionally seafloor surveys have been conducted with research vessels, divers or with an autonomous underwater vehicle (AUV) and are time consuming, expensive and high risk. In this paper we present an approach to merge sonar and monocular images to perform large scale mapping of shallow areas from an autonomous surface vessel (ASV), reducing the mission time, cost and risk. Our method uses multibeam sonar data to generate a mesh of the seafloor. Optical images are then blended and projected onto the mesh after a color correction process which increases contrast and overall image quality. In applicable scenarios, ASVs offer an alternative approach to AUVs for autonomous acoustic and optical site mapping. ASVs are typically less expensive than AUVs and often offer easier deployment and recovery logistics. Also, the mechanical requirements are less demanding because they do not have to withstand increased atmospheric water pressure at depth.

I. INTRODUCTION

Seafloor bathymetry maps are the first requirement for multiple marine and ocean science projects, including fisheries management [1], sediment process modeling, [2] and coral reef monitoring [3]. Traditionally these surveys have been performed by large research vessels equipped with multibeam sonar systems and sidescan sonars as well as towed sonar arrays. In the case of underwater archaeology, the application presented in this paper, surveys have been traditionally performed by divers equipped with measuring tapes, guidelines, and handheld cameras [4]. As this is a very labor intensive task new technological means are being developed. Recently, methods have been proposed where divers utilize a stereoscopic camera system equipped with global positioning system (GPS) and an inertial measurement unit (IMU) [5]. However, this still requires the use of multiple divers and guidelines and is time consuming. To produce complete uniform coverage of a site is challenging as human divers often lack the visibility and overview knowledge to map an area precisely. In contrast,

autonomous underwater vehicles (AUVs) have been used extensively to automate the process of underwater data collection: [6] presents a map of the "Le Lune" shipwreck and [7] and [8] display an overview of the use of AUV for Benthos monitoring, where the authors analyze the use of AUV technology for underwater archeology in the Gulf of Mexico. Operating AUVs in very shallow water is a major challenge as there are difficulties in bottom following due to vehicle draft and minimum depth requirements. Additionally, AUVs are very expensive assets, require expert operators, and have many catastrophic failure modes. Here we propose an alternative for shallow water mapping using an autonomous surface vessel (ASV). They offer a simple and effective way to explore a large shallow area using both optical and acoustic sensors to map the benthos. One advantage in using an ASV to perform mapping is the constant availability of GPS which improves the simplicity and accuracy of platform localization. Deployment of an ASV is simpler, less risky, and less expensive when compared to an AUV.

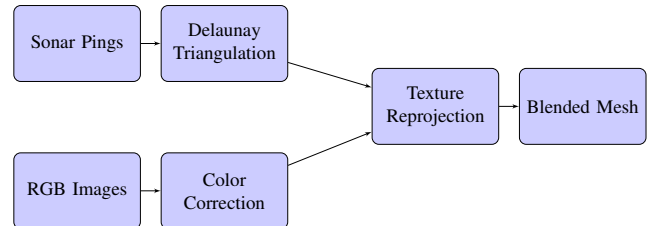


Fig. 1. Flowchart for proposed Sonar and Image processing producing full 3D photo-mosaic from an ASV

This paper describes a method for the use of an ASV for shallow water optical and acoustic 3D mapping. Section II describes the platform used, the sensor payload and general configuration. Section III presents the methodology used for the survey and post-processing, describing both the sonar and camera processing pipelines in detail. Section IV presents the experimental results obtained during a field deployment in April 2015 in Port Royal, Jamaica. Conclusions and future work are presented in Section V.

¹Eduardo Iscar and Matthew Johnson-Roberson are with the Department of Naval Architecture and Marine Engineering, University of Michigan, Ann Arbor, MI 48109 USA {eiscar,mattjr}@umich.edu

Mass	29 kg
Max Payload Mass	10 kg
Size	1.35 x 0.98 x 0.32 m
Max. Speed	3.3 kn
Propulsion System	Differential Waterjet Drive
Autonomy	1 - 2 hrs
Batteries	NiMH 14.4V 20Ah
Onboard Processing	Intel Atom @ 1.8Ghz, 4Gb RAM
GPS receiver	Ublox GPS
Attitude and Heading	CHRobotics UM6 IMU
Sonar	Delta-T Multibeam
Communications	Wifi and Radio
Cameras	2 Greypoint Chamaleon

TABLE I
KINGFISHER ASV MAIN CHARACTERISTICS

II. HARDWARE PLATFORM

- 1) *Kingfisher ASV*: The ASV, from Clearpath Robotics, is a 1.3m long catamaran which has a battery life of approximately 2 hours. The catamaran configuration makes it a very stable vehicle, and offers ample space between both hulls to mount the sensor payload. The Kingfisher is equipped with a UBlox GPS sensor, as well as a low cost IMU whose measurements are integrated into a standard extended Kalman filter (EKF) to determine the vehicle's pose. A full list of the sensor payload appears in Table I. The navigation and control is performed with the Robot Operating System (ROS).
- 2) *Multibeam Sonar*: The vehicle features an Imagenex Delta T multibeam sonar that operates at 260 kHz. It has a swath width of 120°, 480 range bins and a sampling frequency of up to 20 Hz.
- 3) *Camera System*: The camera system is a custom design with two downward facing 1MP Point Grey Chameleon cameras capturing images at 7.5Hz. Image resolution is 1296 x 964 pixels. The cameras are mounted vertically so they face the seafloor and they are positioned close to the boat's center of rotation to reduce the disturbances induced by wave motion.

III. METHODOLOGY

Our proposed solution integrates the measurements collected by the multibeam sonar and RGB camera. We use the 3D structure derived from the multibeam and fuse it with the camera data to produce a textured 3D model for visualization. A flow chart of the process appears in Figure 1. The following section details the steps that are performed:

A. Localization

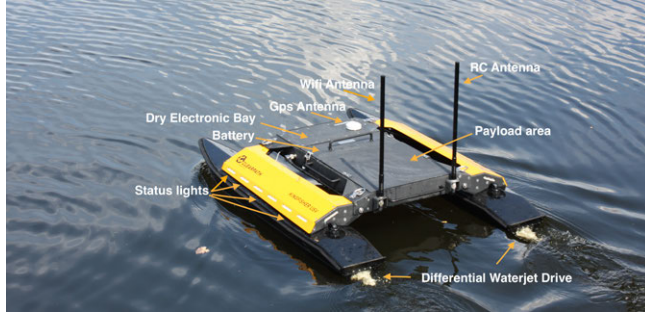
An EKF tracks the position estimate on board the vehicle, fusing IMU and GPS. The availability of GPS and the lack of human-made structures in the vicinity of the vehicle that might interfere with the GPS signal greatly improve the localization accuracy.

B. Sonar data processing

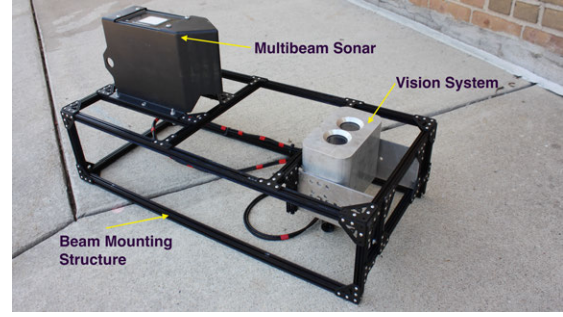
The globally aligned sonar scans are used to estimate a refined roll offset relative to the vehicle frame. This is done by recording two sufficiently large straight segments in opposed directions over an approximately flat seabed. From that data, roll and pitch offsets are computed by obtaining the best fitting planes to each of the directions and calculating the relative slope between the planes. After applying roll and pitch corrections to the data, the resulting sonar pings are median filtered to eliminate outliers. A georeferenced bathymetric grid at a resolution of 15cm is then created from the filtered sonar pings by 2D, thin plate spline interpolation. The last step in the sonar processing pipeline is to produce a triangulated irregular network (TIN) through Delaunay triangulation of the obtained interpolated grid.

C. Image color correction

Images are color corrected to minimize the effect of the water column and lighting pattern using the Greyworld approach [9] to normalize the intensities. This approach treats each pixel color channel independently and constructs an approximate model of the lighting pattern by computing the mean and standard deviation of all channels over all the images captured. A gain and offset is then applied to each channel to increase the distribution contrast. Figures 4(a) and 4(b) show one image before and after applying this process. Another challenge is presented by the different wavelength attenuation of light in water. If the images to be processed have significant depth differences, the illumination correction computed violates the assumption of a uniform path length of light. Without adjusting for depth the single correction applied overamplifies the red channel in the shallow images and the blue channel in the deep images respectively. Figure 4(c) shows an image taken in a shallow part of the data set after applying the correction. To compensate for this we incorporate the depth information from the multibeam sonar and segment the images by depth into bins. The range of each segment depends on the total depth variation in the data set. After segmenting the images, the color correction algorithm is applied to each group independently. Figure 5 illustrates the depth correction

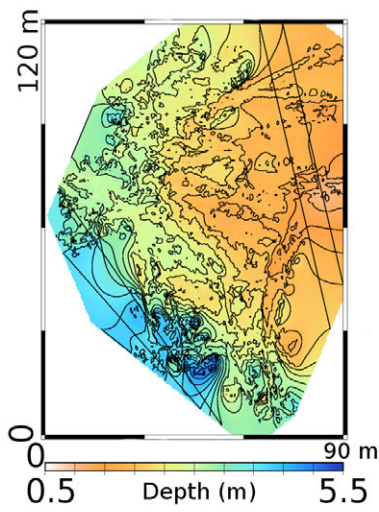


(a) Kingfisher ASV

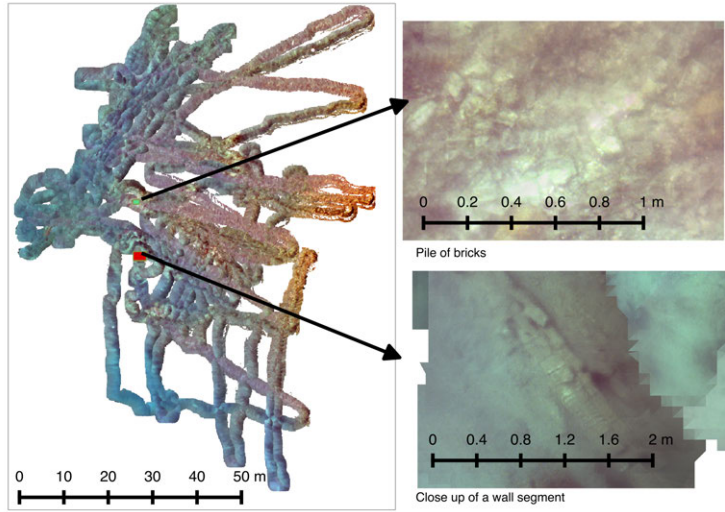


(b) Sensor Payload

Fig. 2. (a) shows the Kingfisher ASV operating, all systems are labeled. (b) shows the sensor payload. Lightweight aluminum beams were used to form a rigid support to attach both the multibeam sonar and the camera vision system to the ASV. The payload is shown upside-down in the figure.



(a) Gridded Sonar Data



(b) Final textured mesh

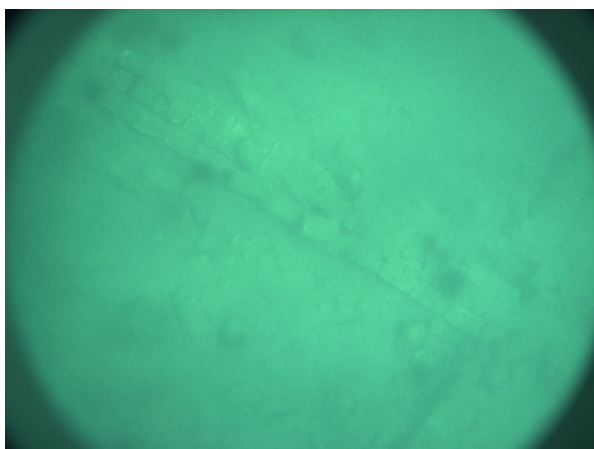
Fig. 3. (a) shows the bathymetry map from the site in Port Royal after being median filtered and gridded. (b) shows the textured, large scale 3D reconstruction of the site. The inset on the top right shows a detailed view of bricks corresponding to a road, while the bottom right inset shows the remains of a building wall. The large area map allows for globally geo-referencing these areas of the site.

steps. The results of the depth segmented correction on the same image as in Figure 4(c) and can be seen in Figure 4(d) after performing this segmentation and correction.

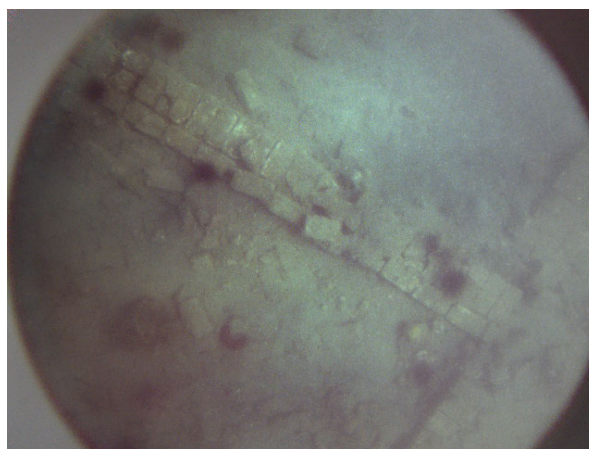
D. Reprojection

The corrected monocular images are then projected and blended onto the mesh using the method proposed by Johnson-Roberson et al. [10], [9]. Using a calibrated camera and the pose of the vehicle the 3D points from sonar mesh can be back-projected into all camera views in which they are visible. The process of determining visibility is accelerated using a KD-Tree [11] making the calculation of image coordinates efficient on large scale models which are hundreds of meters squared. The four closest reprojections are selected to enable a multi-scale band-limited blending

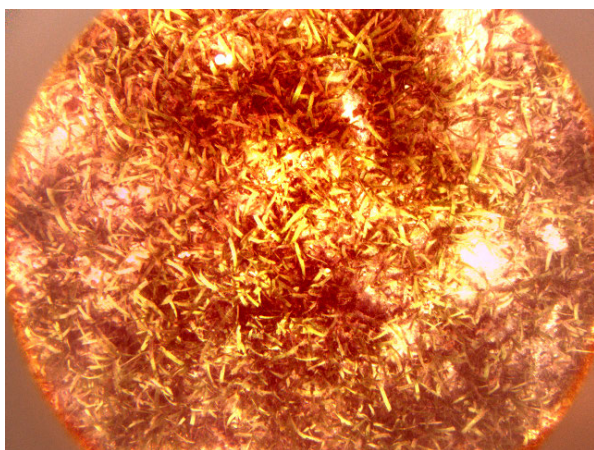
of overlapping images. The initial parametrization is remapped to optimize resolution for the 3D structure in a more hardware efficient square 2D texture space. The resulting mesh parametrization is then segmented to be efficiently packed and stored in a large Virtual texturing pyramid for efficient rendering [12]. The actual creation of the texture images is handled by an out-of-core software renderer detailed in Johnson-Roberson et al. [10]. The sonar derived structure is down-sampled using a Quadric-based method [13] and stored to produce a static level of detail (LOD) hierarchy to allow for the real-time visualization of the model on commodity desktop hardware.



(a) Original image as captured by the camera



(b) Corrected image



(c) Altitude agnostic color correction



(d) Altitude aware color correction

Fig. 4. (a) shows the original image as captured by the camera. (b) shows the same image after applying the color correction algorithm. (c) shows the result of the color correction algorithm if applied to the whole image data set. On the contrary, (d) depicts the same image if images are segmented by depth before processing and then color corrected in batches.

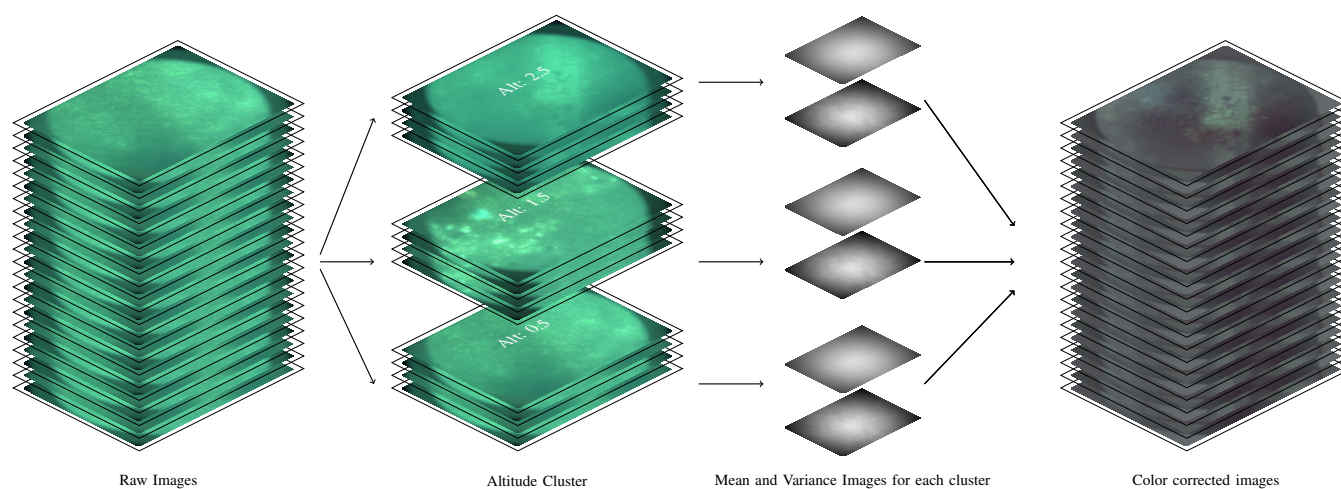


Fig. 5. In this figure the general flow for the attitude aware color correction algorithm is shown: All captured images are clustered into depth bins using the associated sonar measurements. For each cluster, the greyworld assumption is applied independently, computing cluster image mean and variance.

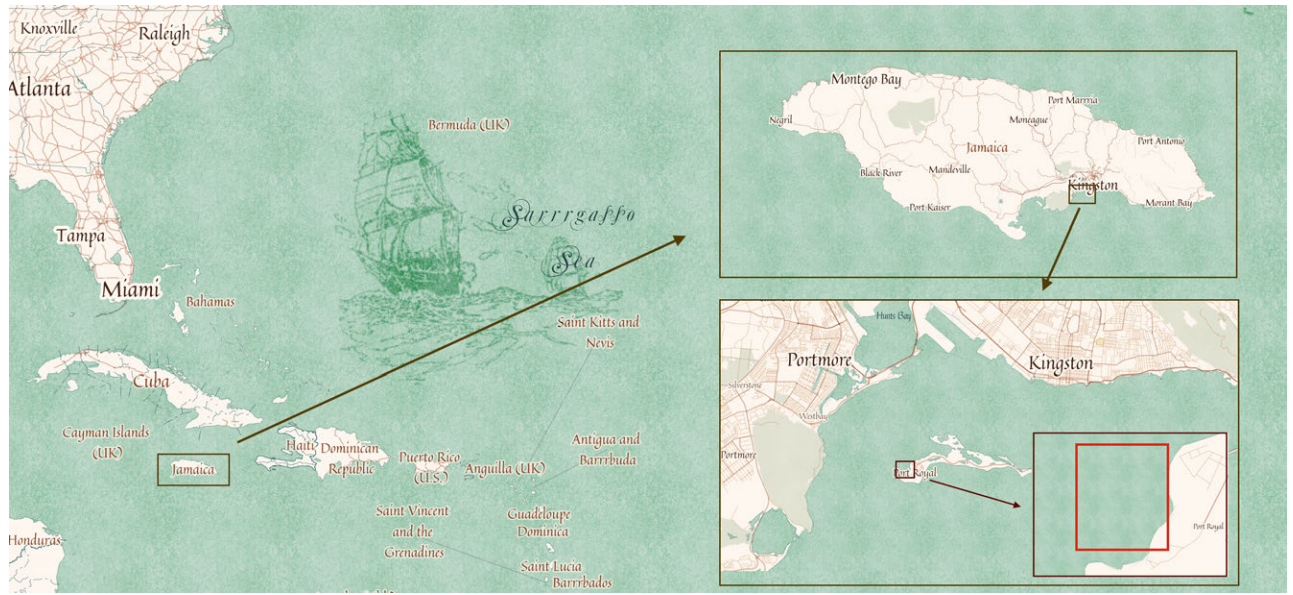


Fig. 6. Map showing the location of the site with inlays showing a higher level of detail of the region of interest. The final mapped area is shown as the red rectangle.

IV. EXPERIMENTS

A. Port Royal

The first deployment of the developed ASV system was April 2015 to map the underwater remains of the city of Port Royal, Jamaica. The exact location of the site mapped can be seen in Figure 6. This site was ideal for such a system due to the shallow depth of the area, with most ruins located between 1 and 3 meters. Archaeologically it is of great importance. The location of Port Royal, in the middle of the Caribbean Sea, and its control over the entrance to Kingston Bay, one of the biggest natural harbors of the region, made it a strategic point of interest for the Europeans, and the city quickly became the center of English trade in the New World. At the same time, it served as port for pirates preying on the Spanish treasure fleets, which gave it its nickname of ‘Wickedest City on Earth’. On June 7th 1692, two thirds of the city disappeared in the waters of Kingston Bay due to an earthquake that had an estimated intensity of 7.5 on the Richter scale. 2000 people died during the earthquake, and 3000 more perished in the following weeks as a consequence of injuries and diseases. With some excavations beginning in the 1960s and 1980s, intact building foundations and streets are perfectly visible in the shallow waters. The site covers an area of approximately 6 acres starting from the coastline and extending about 100m into Kingston Bay.

B. ASV operation

The size of the site makes it extremely difficult to map using conventional means. The draft of most manned vessels would prevent their operation on the site. The small ASV enabled broad coverage of a large area inaccessible by other means. The data presented in Figure 3 displays the results of processing approximately one and a half hours of sensor logs where the vehicle covered more than 2 linear km. In the subfigure on the upper right we can observe bricks from a former street and the remains of a wall appear in the subfigure on the lower right. Problems with the autonomous navigation software made it necessary to teleoperate the ASV. The disadvantages of doing so can be seen in Figure 7, where the trajectory followed by the ASV is shown. A more homogeneous and regular surveying pattern would further increase the quality of the final 3D reconstruction as shown in Figure 3. The large area metric 3D photo-mosaic will enable underwater archaeologists to develop architectural maps of the city and to plan future dives and excavations.

V. CONCLUSIONS AND FUTURE WORK

In this paper we have demonstrated that it is possible to perform surveys and map large areas of shallow seabed with the help of sonar and a monocular camera. The use of a small ASV reduces the time needed to perform such a survey, as well as the cost, and produces accurate, globally referenced three dimensional reconstructions of the seafloor. Future work will focus on improving the quality of

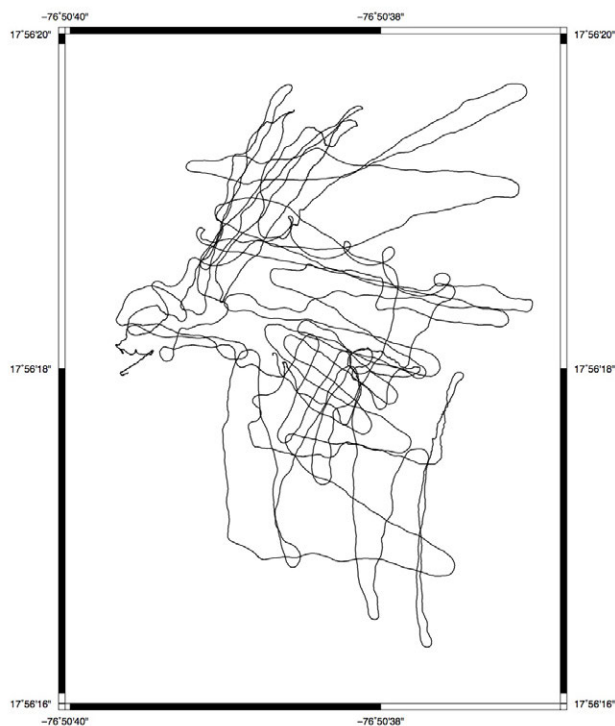


Fig. 7. Map showing the trajectory followed by the Kingfisher ASV. The resulting reconstruction can be seen in Figure 3

the image enhancement process as well as online processing of the images on board the ASV. Other areas of interest are to improve the autonomous navigation capabilities to ensure dense coverage of the area by integrating the sonar depth information to compute the images field of view as the survey is performed.

ACKNOWLEDGMENT

The authors would like to thank the Jamaican National Heritage Trust for the help and support provided, as well as our colleagues Jon Henderson, Andrew Durrant, Katherine Skinner and Jie Li. The work presented has been funded by the National Oceanic and Atmospheric Administration under grant number NA14OAR40110265.

REFERENCES

[1] F. Gerlotto, M. Soria, and P. Fréon, "From two dimensions to three: the use of multibeam sonar for a new approach

[2] P. Dartnell and J. V. Gardner, "Predicting seafloor facies from multibeam bathymetry and backscatter data," *Photogrammetric Engineering & Remote Sensing*, vol. 70, no. 9, pp. 1081–1091, 2004.

[3] J. Roberts, C. Brown, D. Long, and C. Bates, "Acoustic mapping using a multibeam echosounder reveals cold-water coral reefs and surrounding habitats," *Coral Reefs*, vol. 24, no. 4, pp. 654–669, 2005.

[4] A. Bowens, *Underwater Archaeology: The NAS Guide to Principles and Practice*, pp. 114–134. Blackwell Publishing Ltd., 2009.

[5] J. Henderson, O. Pizarro, M. Johnson-Roberson, and I. Mahon, "Mapping submerged archaeological sites using stereo-vision photogrammetry," *International Journal of Nautical Archaeology*, vol. 42, no. 2, pp. 243–256, 2013.

[6] N. Gracias, P. Rida, R. Garcia, J. Escartin, M. L'Hour, F. Cibecchini, R. Campos, M. Carreras, D. Ribas, N. Palomeras, L. Magi, A. Palomer, T. Nicosevici, R. Prados, R. Hegedus, L. Neumann, F. de Filippo, and A. Mallios, "Mapping the moon: Using a lightweight auv to survey the site of the 17th century ship la lune," in *OCEANS - Bergen, 2013 MTS/IEEE*, pp. 1–8, June 2013.

[7] S. B. Williams, O. R. Pizarro, M. V. Jakuba, C. R. Johnson, N. S. Barrett, R. C. Babcock, G. Kendrick, P. D. Steinberg, A. J. Heyward, P. J. Doherty, *et al.*, "Monitoring of benthic reference sites: using an autonomous underwater vehicle," *Robotics & Automation Magazine, IEEE*, vol. 19, no. 1, pp. 73–84, 2012.

[8] D. J. Warren, R. A. Church, K. L. Eslinger, *et al.*, "Deepwater archaeology with autonomous underwater vehicle technology," in *Offshore Technology Conference, Houston, TX*, 2007.

[9] M. Johnson-Roberson, O. Pizarro, S. B. Williams, and I. Mahon, "Generation and Visualization of Large-scale Three-dimensional Reconstructions from Underwater Robotic Surveys," *Journal of Field Robotics*, vol. 27, no. 1, pp. 21–51, 2010.

[10] M. Johnson-Roberson, M. Bryson, B. Douillard, O. Pizarro, and S. B. Williams, "Out-of-core efficient blending for underwater georeferenced textured 3d maps," in *IEEE Computing for Geospatial Research and Application (COM. Geo), 2013 Fourth International Conference on*, pp. 8–15, 2013.

[11] J. L. Bentley, "Multidimensional binary search trees used for associative searching," *Communications of the ACM*, vol. 18, no. 9, pp. 509–517, 1975.

[12] M. Mittring and Crytek GmbH, "Advanced virtual texture topics," in *International Conference on Computer Graphics and Interactive Techniques: Games (SIGGRAPH: Games), SIGGRAPH '08*, pp. 23–51, ACM, 2008.

[13] M. Garland and P. S. Heckbert, "Surface simplification using quadric error metrics," in *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, pp. 209–216, ACM Press/Addison-Wesley Publishing Co., 1997.